This Page Is Inserted by IFW Operations and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.

-- BRIEF DESCRIPTION OF THE DRAWINGS

This invention will be more fully described with reference to the drawings in which:

1

Figure 1 depicts the universal code equivalent of the mitochondrial I-Scel gene (SEQ ID NOS:3 AND 4).

Figure 2 depicts the nucleotide sequence of the invention encoding the enzyme I-Scel and the amino acid sequence of the natural I-Scel enzyme (SEQ ID NOS: 5 and 2).

Figure 3 depicts the I-Scel recognition sequence and indicates possible base mutations in the recognition site and the effect of such mutations on stringency of recognition (SEQ ID NOS: 6, 7, AND 8).

Figure 4A-B is the nucleotide sequence and deduced amino acid sequence of a region of plasmid pSCM525. The nucleotide sequence of the invention encoding the enzyme I-Scel is enclosed in the box (SEQ ID NOS: 9 through 16).

Figure 5 depicts variations around the amino acid sequence of the enzyme I-Scel (SEQ. ID NO:2).

Figure 6 shows Group I intron encoding endonucleases and related endonucleases (SEQ ID NOS: 17 through 44).

Figure 7 depicts yeast expression vectors containing the synthetic gene for I-Scel.

Figure 8 depicts the mammalian expression vector PRSV I-Scel.

Figure 9 is a restriction map of the plasmid pAF100. (See also YEAST, 6:521-534, 1990, which is relied upon and incorporated by reference herein).

Figures 10A and 10B show the nucleotide sequence and restriction sites of regions of the plasmid pAF100 (SEQ ID NOS:45 through 50). --

Figure 11 depicts an insertion vector pTSM ω , pTKM ω , and pTT ω containing the I-Scel site for *E. coli* and other bacteria.

Figure 12 depicts an insertion vector pTYW6 containing the I-Scel site for yeast.

Figure 13A-C depicts an insertion vector PMLV LTR SAPLZ containing the I-Scel site for mammalian cells.

Figure 14A-B depicts a set of seven transgenic yeast strains cleaved by I-Scel. Chromosomes from FY1679 (control) and from seven transgenic yeast strains with I-Scel sites inserted at various positions along chromosome XI were treated with I-Scel. DNA was electrophoresed on 1% agarose (SeaKem) gel in 0.25 X TBE buffer at 130 V and 12°C on a Rotaphor apparatus (Biometra) for 70 hrs using 100 sec to 40 sec decreasing pulse times. (A) DNA was stained with ethidium bromide (0.2 μg/ml) and transferred to a Hybond N (Amersham) membrane for hybridization. (B)³²P labelled cosmid pUKG040 which hybridizes with the shortest fragment of the set was used as a probe. Positions of chromosome XI and shorter chromosomes are indicated.

Figure 15A-E depicts the rationale of the nested chromosomal fragmentation strategy for genetic mapping. (A) Positions of *I-Sce*l sites are placed on the map, irrespective of the left/right orientation (shorter fragments are arbitrarily placed on the left). Fragment sizes as measured from PFGE (Fig. 14A) are indicated in kb (note that the sum of the two fragment sizes varies slightly due to the limit of precision of each measurement). (B) Hybridization with the probe that hybridizes the shortest fragment of the set determines the orientation of each fragment (see Fig. 14B). Fragments that hybridize with the probe (full lines) have been placed arbitrarily to the left. (C) Transgenic yeast strains have been ordered with increasing sizes of hybridizing chromosome fragments. (D) Deduced *I-Sce*l map with minimal and maximal size of intervals indicated in kb (variations in some intervals are due to limitations of PFGE measurements). (E) Chromosome subfragments are used as probes to assign each cosmid clone to a given map interval or across a given *I-Sce*l site.

Figure 16A-C depicts mapping of the *I-Sce*I sites of transgenic yeast strains by hybridization with left end and right end probes of chromosome XI. Chromosomes from FY1679 (control) and the seven transgenic yeast strains were digested with *I-Sce*I. Transgenic strains were placed in order as explained in Fig. 15. Electrophoresis conditions were as in Fig. 14. ³²P labelled cosmids pUKG040 and pUKG066 were used as left end and right end probes, respectively.

Figure 17A-H depicts mapping of a cosmid collection using the nested chromosomal fragments as probes. Cosmid DNAs were digested with EcoRI and electrophoresed on 0.9% agarose (SeaKem) gel at 1.5 V/cm for 14 hrs, stained with ethidium bromide and transferred to a Hybond N membrane. Cosmids were placed in order from previous hybridizations to help visualize the strategy. Hybridizations were carried out serially on three identical membranes using left end nested chromosome fragments purified on PFGE (see Fig. 16) as probes. A: ethidium bromide staining (ladder is the BRL "1kb ladder"), B: membrane #1, probe: Left tel to A302 site,

C: membrane #1, probe: Left tel to M57 site, D: membrane #2, probe: Left tel to H81 site, E: membrane #2, probe: Left tel to T62 site, F: membrane #3, probe: Left tel to G41 site, G: membrane #3, probe: Left tel to D304 site, H: membrane #3, probe: entire chromosome XI.

Figure 18 depicts a map of the yeast chromosome XI as determined from the nested chromosomal fragmentation strategy. The chromosome is divided into eight intervals (with sizes indicated in kb, see Fig. 15D) separated by seven *I-Sce*I sites (E40, A302 ...). Cosmid clones falling either within intervals or across a given *I-Sce*I site are listed below intervals or below interval boundaries, respectively. Cosmid clones that hybridize with selected genes used as probes are indicated by letters (a-i). They localize the gene with respect to the *I-Sce*I map and allow comparison with the genetic map (top).

Figure 19A-B depicts diagrams of successful site directed homologous recombination experiments performed in yeast.

Figure 20A-E. Experimental design for the detection of HR homologous recombination (HR) induced by I-Sce I. a) Maps of the 7.5 kb *tk -PhleoLacZ* retrovirus (G-MtkPL) and of the 6.0 kb *PhleoLacZ* retrovirus (G-MPL), SA is splice acceptor site. G-MtkPL sequences (from G-MtkPL virus) contains *PhleoLacZ* fusion gene for positive selection of infected cells (in phleomycin-containing medium) and *tk* gene for negative selection (in gancyclovir-containing medium). G-MPL sequences (from G-MPL virus) contains only *PhleoLacZ* sequences. b) Maps of proviral structures following retroviral integration of G-MtkPL and G-MPL. I-Sce I *PhleoLacZ* LTR duplicates, placing I-Sce I *PhleoLacZ* sequences in the 5' LTR. The virus vector (which functions as a promoter trap) is transcribed (arrow) by a flanking cellular promoter, P. c) I-Sce I creates two double strand breaks (DSBs) in host DNA liberating the central segment and leaving broken chromosome ends that can pair with the donor plasmid, pVRneo (d). e) Expected recombinant locus following HR.

Figure 21<u>A-B.</u> **A.** Scheme of pG-MPL. SD and SA are splice donor and splice acceptor sites. The structure of the unspliced 5.8 kb (genomic) and spliced 4;2 kb transcripts is shown below. Heavy bar is 32 P radiolabelled LacZ probe (**P**). B. RNA Northern blot analysis of a pG MLP transformed ψ -2 producer clone using polyadenylated RNA. Note that the genomic and the spliced mRNA are produced at the same high level.

Figure 22A-B. A. Introduction of duplicated I-Sce I recognition sites into the genome of mammalian cells by retrovirus integration. Scheme of G-MPL and G-MtkPL proviruses which illustrates positions of the two LTRs and pertinent restriction sites. The size of *Bcl* I fragments and of I-Sce I fragments are indicated. Heavy bar is ³²P radiolabelled *LacZ* probe (**P**). **B.** Southern blot analysis of cellular DNA from NIH3T3

fibroblasts cells infected by G-MtkPL and PCC7-S multipotent cells infected by G-MPL. Bcl I digests demonstrating LTR mediated PhleoLacZ duplication; I-Sce I digests demonstrating faithful duplication of I-Sce I sites.

Figure 23A-B. Verification of recombination by Southern. A.: Expected fragment sizes in kilobase pairs (kb) of provirus at the recombinant locus. 1) the parental proviral locus. Heavy bar (P) is ³²P radioactively labelled probe used for hybridization. 2) a recombinant derived after cleavage at the two I-Sce I sites followed by gap repair using pVR neo (double-site homologous recombination, DsHR). 3) a recombination event initiated by the cleavage at the I-Sce I sites in the left LTR (single-site homologous recombination, SsHR). B.: Southern analysis of DNA from NIH3T3/G-MtkPL clones 1 and 2, PCC7-S/G-MPL clones 3 and 4 and transformants derived from cotransfection with pCMV(I-Sce I+) and pVRneo (1a, 1b, 2a, 3a, 3b and 4a). *Kpn* I digestion of the parental DNA generates a 4.2 kb fragment containing *LacZ* fragment. Recombinants 1a and 3a are examples of DsHR Recombinants 1b, 2a, 3b and 4a are examples of SsHR.

Figure 24A-B. Verification of recombination by Northern blot analyses. A.: Expected structure and sizes (in kb) of RNA from PCC7-S/G-MPL clone 3 cells before (top) and after (bottom) I-Sce I induced HR with pVRneo.I Heavy bars P1 and P2 are ³²P radioactively labelled probes. B.: Northern blot analysis of the PCC7-S/G-MPL clone 3 recombinant (total RNA). Lane 3 is parental cells, lane 3a recombinant cells. Two first lanes were probed with *LacZ* P1, two last lanes are probed with *neo* P2. parental PCC7-S/G-MPL clone 3 cells express a 7.0 kb *LacZ* RNA as expected of trapping of a cellular promoter leading to expression of a cellular-viral fusion RNA. The recombinant clone does not express this *Lacz* RNA but expresses a *neo* RNA of 5.0 kb, corresponding to the size expected for an accurate replacement of *PhleoLacZ* by *neo* gene.

Figure 25A-C. Types of recombination events induced by I-Sce I DSBs, a) Schematic drawing of the structure of the recombination substrate. The G-MtkPL has provirus two LTRs, each containing an I-Sce I recognition site and a *PhleoLacZ* gene. The LTRs are separated by viral sequences containing the *tk* gene. The phenotype of G-MtkPL containing cells is Phleo^R, GIs^s, β-Gal± b) Possible modes of intra-chromosomal recombination. 1) The I-Sce I endonuclease cuts the I-Sce I site in the 5'LTR. The 5' part of U3 of the 5'LTR can pair and recombine with it homologous sequence in the 3'LTR (by single-strand annealing (SSA). 2) The I-Sce I endonuclease cuts the I-Sce I site in the 3'LTR. The 3' part of U3 of the 3'LTR can pair and recombine with its homologous sequence in the 5'LTR (by SSA). 3) The I-Sce I endonuclease cuts I-Sce I sites in the two LTRs. The two free ends can relegate (by an end-joining mechanism). The resulting recombination product in each of the three models is a solitary LTR (see right side). No modification would occur in the cellular sequences flanking the integration site. c) The I-Sce I endonuclease cuts the I-Sce I sites in the

two LTRs. The two free ends can be repaired (by a gap repair mechanism) using the homologous chromosome. On the right, the resulting recombination product is the deletion of the proviral integration locus.

- Figure 26A-C. Southern blot analysis of DNA from NIH3T3/G-MtkPL 1 and 2, and *PhleoLacZ* recombinants derived from transfections with pCMV(I-Sce I+) selected in Gancyclovir containing medium. a) Expected fragment sizes in kilobase pair (kbp) of parental provirus after digestion with *Pst* I endonuclease. *Pst* I digestion of the parental DNA NH3T3/G-MtkPL 1 generates two fragments of 10 kbp and of the parental NIH3T3/G-MtkPL 2 two fragments of 7 kbp and 9 kbp. b) Southern blot analysis of DNA digested by *Pst* I from NIH3T3/G-MtkPL 1, and recombinants derived from transfection with pCMV(I-Sce I+) (1.1 to 1.5). c) Southern blot analysis of DNA digested by *Pst* I from NIH3T3/G-MtkPL 2, and recombinants derived from transfection with pCMV(I-Sce I+) (2.1 to 2.6). Heavy bar is ³²P radiolabelled *LacZ* probe (P).
- Figure 27A-C. Southern blot analysis of DNA from NIH3T3/G-MtkPL 1 and 2, and *PhleoLacZ*⁺ recombinants derived from transfections with pCMV(I-Sce I+) and pCMV(I-Sce I-) and selection in Phleomycin and Gancyclovir containing medium.

 a1) Expected fragment sizes in kbp of parental provirus after digestion with *Pst* I or *Bcl* I endonuclease. *Pst* I digestion of the parental DNA NIH3T3/G-MtkPL 1 generates two fragments of 10 kbp. *Bcl* I digestion of the parental DNA NIH3T3/G-MtkPL 2 generates three fragments of 9.2 kbp, 7.2 kbp and 6.0 kbp. a2) Expected fragment sizes in kbp of recombinants after digestion with *Pst* I or *Bcl* I endonuclease. *Pst* I digestion of DNA of the recombinant derived from NIH3T3/G-MtkPL 1 generates one fragment of 13.6 kbp. *Bcl* I digestion of the DNA of the recombinants derived from NIH3T3/G-MtkPL 2 generates two fragments of 9.2 kbp and 6.0 kbp. b) Southern blot analysis of DNA from NIH3T3/G-MtkPL 1, and recombinants derived from transfection with pCMV(I-Sce I-) and pCMV(I-Sce I+) (1c, 1d). c) Southern analysis of DNA from NIH3T3/G-MtkPL 2, and transformants derived from transfection with pCMV(I-Sce I-) (2a, 2b) and pCMV(I-Sce I+) (2c to 2h). Heavy bar is ³²P radiolabelled *LacZ* probe (P).
- **Figure 28.** Figure 28 is a diagram illustrating the loss of heterozygosity by the insertion or presence of an I-Sce I site, expression of the enzyme I-Sce I, cleavage at the site, and repair of the double strand break at the site with the corresponding chromatid.
- **Figure 29.** Figure 29 is a diagram illustrating conditional activation of a gene. An I-Sce I site is integrated between tandem repeats, and the enzyme I-Sce I is expressed. The enzyme cleaves the double stranded DNA at the I-Sce I site. The double strand break is repaired by single stand annealing, yielding an active gene.
- **Figure 30.** Figure 30 is a diagram illustrating one step rearrangement of a gene by integration of an I-Sce I site or by use of an I-Sce I site present in the gene. A

plasmid having either one I-Sce I site within an inactive gene, or two I-Sce I sites at either end of an active gene without a promoter, is introduced into the cell. The cell contains an inactive form of the corresponding gene. The enzyme I-Sce I cuts the plasmid at the I-Sce I sites, and recombination between the chromosome and the plasmid yields an active gene replacing the inactive gene.

- **Figure 31.** Figure 31 is a diagram illustrating the duplication of a locus. An I-Sce I site and a distal part of the locus are inserted into the gene by classical gene replacement. The I-Sce I site is cleaved by I-Sce I enzyme, and the break is repaired by homologous sequences. This results in duplication of the entire locus.
- **Figure 32.** Figure 30 32 is a diagram illustrating the deletion of a locus. Two I-Sce I sites are added to flank the locus to be deleted. The I-Sce I enzyme is expressed, and the sites are cleaved. The two remaining ends recombine, deleting the locus between the two I-Sce I sites.
- **Figure 33.** Figure 33 is a diagram of plasmid pG-MtkΔPAPL showing the restriction sites. The plasmid is constructed by deletion of the polyadenylation region of the tk gene from the pGMtkPL plasmid. --

On page 18, replace the last paragraph with the following amended paragraph:

-- The enzyme I-Scel has a known recognition site. (ref. 14.) The recognition site of I-Scel is a non-symmetrical sequence that extends over 18 bp as determined by systematic mutational analysis. The sequence reads: (arrows indicate cuts)

On pages 47-48, replace the paragraph bridging these pages with the following amended paragraph:

-- -e- The supernatant of this clone was used to infect other mouse cells (1009) by spreading 10⁵ virus particles on 10⁵ cells in DMEM medium with 10% fetal calf serum and 5 mg/ml of "polybrain" polybrene (hexadimethrine bromide). Medium was replaced 6 hours after infection by the same fresh medium. --

On pages 71-72, replace the paragraph bridging these pages with the following amended paragraph:

-- This example describes the use of the I-Sce I meganuclease (involved in intron homing of mitochondria of the yeast Saccharomyces cerevisiae) (6B, 28B) to induce DSB and mediate recombination in mammalian cells. I-Sce I is a very rare-cutting restriction endonuclease, with an 18 bp long recognition site (29B, 22B). In vivo, I-Sce I endonuclease can induce recombination in a modified yeast nucleus by initiating a specific DBS leading to gap repair by the cell (30B, 17B, 21B). Therefore, this approach can potentially be used as a means of introducing specific DSB in chromosomal target DNA with a view to manipulate chromosomes in living cells. The I-Sce I-mediated recombination is superior to recombinase system [11B] for chromosome engineering since the latter requires the presence of target sites on both host and donor DNA molecules, leading to reaction that is reversible. --

On pages 72-73, replace the paragraph bridging these pages with the following amended paragraph:

-- pG-MtkPL was obtained in five steps: (I) insertion of the 0.3 kbp *Bgl II-Sma* I fragment (treated with Klenow enzyme) of the Moloney Murine Leukemia Virus (MoMuLV) *env* gene (25B) containing a splice acceptor (SA) between the *Nhe* I and *Xba* I sites (treated with Klenow enzyme), in the U3 sequence of the 3'LTR of MoMuLV, in an intermediate plasmid. (II) Insertion in this modified LTR of a 3.5 kbp *Nco* I-*Xho* I fragment containing the *PhleoLacZ* fusion gene [13B] (from pUT65; Cayla Laboratory, Zone Commerciale du Gros, Toulouse, France) at the *Xba* I site next to SA. (III) Insertion of this 3'LTR (containing SA and *PhleoLacZ*), recovered by *Sal* I-*EcoR* I double digestion in the p5'LTR plasmid (a plasmid containing the 5'LTR up to the nucleotide no 563 of MoMuLV [12B]) between the *Xho* I and the *EcoR* I site. (IV) Insertion of a *synthetic* I-*Sce* I recognition site into the *Nco* I site in the 3'LTR (between SA and *PhleoLacZ*), and (V) insertion (antisense to the retroviral genome) of the 1.6 kbp *tk* gene with its promoter with linker adaptators at the *Pst* I site of pG-MPL. --

On page 80, replace the paragraph beginning on line 5 with the following amended paragraph:

-- The generation of *tk-lPhleoLacZ* cells is probably a consequence of either a homo-allelic and/or an ectopic gene conversion event (36B). Isolation and detailed molecular analysis of the proviral integration sites will provide information on the relative frequency of each of these events for the resolution of chromosomal DSBs by the cell. This quantitative information is important as, in mammalian cells, the high redundancy

of genomic sequences raises the possibility of a repair of DSBs by ectopic homologous sequences. Ectopic recombination for repair of DSBs may be involved in genome shaping and diversity in evolution [29]. --

After making the above amendments, please move pages 50-58 and insert these pages after page 88 and before page 89 of the application.

Renumber pages 50-58 as pages 80-88, respectively.

Renumber pages 59-88 as pages 50-79, respectively.

On pages 89-93, replace the section entitled "References" with the following amended section.

-- References

- 1 B. Bernstein, N., Pennell, N., Ottaway, C. A. and Shulman, M.J.
- 1992. Gene replacement with one-sided homologous recombination. Mol. Cell Biol. 12: 360-367.
- 2 <u>B</u>. Bonnerot, C., Legouy, E., Choulika, A. and Nicolas, J.-F. 1992. Capture of a cellular transcriptional unit by retrovirus: mode of provirus activation in embryonal carcinoma cells. J. Virol. 66: 4982-4991.
- 3 <u>B</u>. Bonnerot, C., and Nicolas, J.-F. 1993. Application of LacZ gene fusions to post-implantation development. In "Methods in Enzymology: Guide to techniques in mouse development". 451-469. Wassarman, P.M., DePamphilis, M.L.
- 4 <u>B</u>. Brenner, D.A., Smogocki, A. and Camerini-Otero, R.D. 1986. Double-strand gap repair results in homologous recombination in mouse L cells. Proc. Natl. Acad. Sci. USA. 83: 1762-1766.
- 5 <u>B</u>. Capecchi, M.R. 1989. Altering the genome by homologous recombination. Science. 244: 1288-1292.
- 6 <u>B</u>. Jacquier, A. and Dujon, B. 1985. An intron encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell. 41: 383-394.

- 7 <u>B</u>. Jakob, H. and Nicolas, J.F. 1987. Mouse Tertocarcinoma Cells. In "Methods in Enzymology: Cell lines for genetic analysis". 66-81. Gottesman, M.H., ed., Academic Press.
- 8 <u>B</u>. Jessberger, R. and Berg, P. 1991. Repair of deletions and double-strand gaps by homologous recombination in a mammalian in vitro system. Mol. Cell Biol. 11: 445-457.
- 9 <u>B</u>. Kilby, N.J., Snaith, M.R. and Murray, J.A.H. 1993. Site-specific recombinases: tools for genome engineering. Reviews. 9: 413-421.
- 10 <u>B</u>. Lin, F.L.M., Sperle, K. and Sternberg N. 1990. Repair of double-stranded DNA breaks by homologous DNA fragments during transfer of DNA into mouse L cells. Mol. Cell Biol. 10:113-119.
- 11 <u>B</u>. Lin, F.L.M., Sperle, K. and Sternberg N. 1990. Intermolecular recombination between DNAs introduced into mouse L cells is mediated by a nonconservative pathway that leads to crossover products. Mol. Cell Biol. 10: 103-112.
- 12 <u>B</u>. Lin, F.L.M., Sperle, K. and Sternberg N. 1990. Intermolecular recombination between DNAs introduced into mouse L cells is mediated by a nonconservative pathway that leads to crossover products. Mol. Cell. Biol. 10: 103-112.
- 13 <u>B</u>. Mann, R., Mulligan, R.C. and Baltimore, D. 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell. 33: 153-160.
- 14 <u>B</u>. Mansour, S.L., Thomas K.R. and Capecchi, M.R. 1988. Disruption of the proto-oncogène int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable gene. Nature. 336:348-352.
- 15 <u>B</u>. Mulsant, P., Gatignol, A., Dolens, M. and Tiraby, G. 1988. Somat. Cell. Mol. Genet. 14: 243-252.
- 16 <u>B</u>. Nicolas, J. F. and Rubenstein, J. 1987. Retroviral vectors. Boston London Durban Singapore Sydney Toronto Wellington, Butterworths.

- 17 <u>B</u>. Plessis, A., Perrin, A., Haber, J.E. and Dujon, B. 1992. Site specific recombination determined by I-Sce I, a mitochondrial group I intron-encoded endonuclease expressed in the yeast nucleus. Genetics 130:451-460.
- 18 <u>B</u>. Sauer, B. and Henderson, N. 1988. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Prac. Natl. Acad. Sci. USA. 85: 5166-5170.
- 19 <u>B</u>. Seidman, M.M. 1987. Intermolecular homologous recombination between transfected sequences in mammalian cells is primarily nonconservative. Mol. Cell. Biol. 7: 3561-3565.
- 20 <u>B</u>. Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A. and Kucherlapati, R.S. 1985. Insertion of DNA sequences into the human chromosomal B-globin locus by homologous recombination. Nature. 317: 230-234.
- 21 <u>B</u>. Szostak, J. W., Orr-Weaver, T.L. and Rothstein, R.J. 1983. The double-strand break repair model for recombination. Cell. 33: 25-35.
- 22 <u>B</u>. Thierry, A., Perrin, A., Boyer, J., Fairhead, C., Dujon, B., Frey, B. and Schmitz, G. 1991. Cleavage of yeast and bacteriophage T7 genomes at a single site using the rare cutter endonuclease I-Sce I. Nucleic Acids Res. 19: 189-90.
- 23 <u>B</u>. Tybulewicz, V.L. J., Crawford, C.E., Jackson, P.K., Bronson, R.T. and Mulligan, R.C. 1991. Neonatal Lethality and Lymphopenia in Mice with a Homozygous Disruption of the c-abl Proto-Oncogene. Cell 65: 1153-1163.
- 24 <u>B</u>. Varmus, H. and Brown, P. 1989. Retroviruses <u>in Mobile DNA, 58-108 (Douglas E. Berg and Martha H. Howe eds).</u>
- 25 <u>B</u>. Weiss, R., Teich, N., Varmus, H. and Coffin, J. 1985. RNA tumor viruses. Molecular Biology of tumor viruses. Second Edition. 2) Supplements and appendixes. Cold Spring Harbor Laboratory. 1-1222.
- 26 <u>B</u>. Weiss, R., Teich, N., Varmus, H. and Coffin, J. 1985. RNA tumor viruses. Molecular Biology of tumor viruses. Second Edition. 2) Supplements and appendixes. Cold Spring Harbor Laboratory. 1-1222.

- 27 <u>B</u>. Phillips J. and Morgan W. 1994. Illegitimate recombination induced by DNA double-strand breaks in mammalian chromosomes. Molecular and Cellular Biology 14:5794-5803.
- 28 <u>B</u>. Dujon B. 1989. Group I introns are mobile genetic elements: facts and mechanistic speculations-a review. Gene 82:91-114.
- 29 <u>B</u>. Colleaux L., D'Aurio L., Galibert F. and Dujon B. 1988. Recognition and cleavage site of the intron-encoded omega transposase. Proc Natl Acad Sci USA 85:6022-6.
- 30 <u>B</u>. Fairchild C. and Dujon B. Consequences of unique double-stranded breaks in yeast chromosomes: death or homozygosis. Molecular general genetics 240:170-180.
- 31 <u>B</u>. Pfeiffer P., Thode S., Hancke J. and Vielmetter W. 1994. Mechanism of overlap information in nonhomologous DNA end joining. Molecular and Cellular Biology 14:888-895.
- 32 <u>B</u>. Mezard C. and Nicholas A. 1994. Homologous, homeologous, and illegitimate repair of double-strand breaks during transformation of a wild-type strain and a rad52 Mutant strain of Saccharomyces cerevisiae. Molecular and Cellular Biology 14:1278-1292.
- 33 <u>B</u>. Feaver W.J., Svejstrup J.Q., Bradwell L., Bradwell A.J., Buratowski S., Gulyas K., Donahue T.F., Friedberg E.C. and Kornberg R.D. 1993. Dual Roles of a Multiprotein Complex from S. cerevisiae in transcription and DNA Repair. Cell 75:1379-1387.
- 34 <u>B</u>. Kramer K., Brock J., Bloom K., Moore K. and Haber J. 1994. Two different types of double-strand breaks in Saccharomyces ceerevisiae are repaired by similar RAD52 independent, nonhomolgous recombination events. Molecular and Cellular Biology 14:1293-1301.
- 35 <u>B</u>. Weiffenbach B. and Haber J. 1981. Homothallelic mating type switching generates lethal chromosomes breaks in rad52 strains of Saccharomyces cerevisiae. Molecular and Cellular Biology 1:522-534.

- 36 <u>B</u>. Nassif N., Penney J., Pal S., Engels W. and Gloor G. 1994. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. Molecular and cellular biology 14:1643-1625.
- 37 <u>B</u>. Charlesworth B., Sniegowski P. and Stephan W. 1994. The evolutionary dynamics of repetitive DNA in eucaryotes. Nature 371:215-220. --

After page 93, and before page 94, insert the attached pages titled "SEQUENCE LISTING".

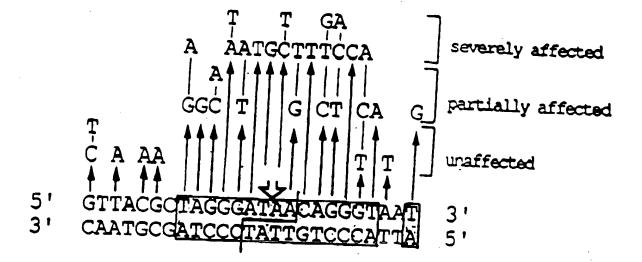
The Universal Code Equivalent of the Mitochondrial 1-Sce I Gene.

The synthetic I-Sce I gene

Born 41

CCGGATCCATE CAT ATG AAA AAC ATC AAA AAA AAC CAG GTA ATG AAC CTG GET CCG AAC TCT AAA CTG CTG AAA GAA TAC AAA TCC CAG CTG ATC GAA CTG AAC ATC GAA CAG TTC GAA SCA GGT ATC GGT CTG ATC CTG GGT GAT GGT TAC ATC CGT TCT CGT GAT GAA GET AAA ACC TAC TGT ATG CAG TTC GAG TER AAA AAC AAA SCA TAC ATG GAC CAC GTA TGT CTG CTG TAC GAT 'K K CAG TEG GTA CTO TEC CEG CEE CAC AMA AMA CAA COT GTT AME CAC CTO GOT AME CTO GTA ATC ACC THE BOC GCC CAR ACT TTC AAA CAC CAA BCT TTC AAC AAA CTE GCT AAC CTG TTC ATC GTT AAC AAC AAA AAA ACC ATC CCB AAC AAC CTB GIT GAA AAC TAC CTB ACC CCG AIG TET CTG GCA TAC TGG TIC ATG GAT GAT GGT AAA TGG GAT TAC AAC AAA AAC TET ACC М AAC AAA TEG ATE GTA CTG AAC ACE CAR TET TTE ACT TTE GAA GAA GTA GAA TAE CTG GTT ANG GET CTG CGT AND AMA TTC CAM CTG MAC TGT TAC GTA AMA ATC MAC AMA CCG ATC ATC TAC ATC BAT TOT ATG TOT TAC CTB ATC TAC AAC CTE ATC AAA CCB TAC CTE Y ATC CCE CAG ATE ATE TAC AAA CTG CCE AAC ACT ATC TCC TCU GAA ACT TIC CTG AAA TAA TAAGTCGACTGCAGGATCCGGTAAGTAA Sail Patl Bank!

1 and 2: These amino acids are absolutely necessary to produce catalytic activity. Other substitutions are possible, such as deletions of the 10 first amino acids.



1447 GEREALMONTSTRUCCUSCUSTUMCACCOCCAMA CONCACA CONCACA AND CONCERNING THE CONCACA AND SOCIAL AND 1811 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1																	•							
1819 STATEMENTS CONCINCTION TO ACCUSE OF THE ARM AND COTE OF ARM SOL ALL ALL ALL ALL ALL ALL ALL ALL ALL A	16	67	α	عد	NOTE:) ICC	271	ras	عديم	œia	ECAN.	يصد		COCN	~44~	SCA CA	r e e							
1819 CET ME SAA ANA COC CAS CAA COC CAS CAT THE NOT OFF COS CAS COT CAS COT 14' 1 1 2 2 2 1 1 2 2 2 1 1 2 2 2 1 1 2 2 2 1 1 2 2 2 1 1 2 2 2 2 1 1 2 2 2 2 1 1 2	170	47	111	124	recti	erca	- A		- 9 ~ -	~~ .								- 46-6	المتحد	west	= = ;	:txtc:	1.144	
1819 CET ME SAA ANA COC CAS CAA COC CAS CAT THE NOT OFF COS CAS COT CAS COT 14' 1 1 2 2 2 1 1 2 2 2 1 1 2 2 2 1 1 2 2 2 1 1 2 2 2 1 1 2 2 2 2 1 1 2 2 2 2 1 1 2	•	1				-,		••••	_,	• • • • • • • • • • • • • • • • • • • •	#6. T 14	-	-100	1117	ttet.	: X5	ن: •	5 570	. Act	344	ند ن	نعد ،	1818	
1819 THE WEAR CAT GTT CTT TOT 1960 GTT ATC CTC TOA ITCHTGRATAACGGTA-MACHGCTTHATGATTAACGGTA- 1818 THATAGCTATULUTGAGGGAGGAGGGAGGGAGGGAGGAGGAGGAGGAGGAGGA																-	•	•		•		•	•	
1819 THE WEAR CAT GTT CTT TOT 1960 GTT ATC CTC TOA ITCHTGRATAACGGTA-MACHGCTTHATGATTAACGGTA- 1818 THATAGCTATULUTGAGGGAGGAGGGAGGGAGGGAGGAGGAGGAGGAGGAGGA		ï	7	H		~ K	n CG	د ص	5 52	A CC	: 34X			10	. 1		: œ	: ::	-	: ::	ac c	; ; ; ;	1419	
1948 TOATMOCRITURIZEDAGECGRACGCCACCCCACCCCACCCCACCCCCACCCCCACCCCCCACCCC								_	-		_	_	•	•	•	-	•	-	Ļ	Ŀ	λ	•	, 1	
1948 TOATMOCRITURIZEDAGECGRACGCCACCCCACCCCACCCCACCCCCACCCCCACCCCCCACCCC		7 7	TV.			r GT	ı ÇT		: ica	: GI	, VIC	: cc	: ::	1 177	7670	CATI	ucc	71-	· *	٠	7646	Texas		
1986 COC CAME TON SCIENCE CAT CAT CAT CAME CAME CAME CAME CAME CAME CAME CAME					•.						•	•											3.0	
1986 COC CAME TON SCIENCE CAT CAT CAT CAME CAME CAME CAME CAME CAME CAME CAME	194	1.6		I ACC	75.75	W.	ELAG	ccev	N CGX(CA.	χœ	600	الكاري	vara	LECEA	مص	œ	uci	ææ		-1~			
1984 COC CHA TRA SCHEMANICATIA ATE IGA GILAGETRACTICACTUTA TAGGORITACACTUTA ATG 1164 2167 2168 2169 2169 2169 2169 217 2169 217 217 217 217 217 217 217 21	202	2 0	coc	ETC	·	: acc./	TIN	-				1-	_							•	• • • • • • • • • • • • • • • • • • • •		7-1	
1988 COC CRG TRA SCHEAAGGAATTA ATG TOM STEATCHTCATTEATHCACCTT AGGSTTACACTT ATG 1164 117 117 116 116 117		1								,	` 7 .	3	L	, 24	l CGA	CAG	G T		းသ	् रा ट			· · · · ·	
2144 AT COC GOT COT AND COT GAT AND COT AND CO	201	34	cae			-										. •	-			•	_	-		
2129 ACC ATC APC OT COT GAT ACT TO THE ACC APC TATO AND ACT TO THE ACCOUNT ATC ATC TO THE ACCOUNT ATC ACC ATC APC APC APC APC APC APC APC APC APC AP	1	3	c	9	•	• •		~~~	1111	K	IWA	GITA	حت	ACTO	YIIF	æς λ	ccα	AGGC	71:X	CACT			::64	
2229 MCC ANG APP AND ARE TOT GAT GOT TOA CASCETATIONTON AND GOD ON AND TOTALICAC 1195 2 TH I THE SERVICE OF TOA CASCETATION AND COSC ON AN	214		~~	~																			•	
2229 ACC ANG Nº 3/03 ANE NOT CAR CITE TOA CAMOCHATCATCANTAMGETTTA AND CHE FAG MITATCAC 2255 2 T H I T P R S R U T T A CAMOCHATCATCANTAMGETTTA AND CHE FAG MITATCAC 2255 2296 ACCTAMATIGCTAACGCACTCAGGCACTCTTC ATG AND TOT AND COG CTC ATC STC ATC CTC GGC 2253 2316 ACC GTC ACC CTG CAT GCT GTA GGC ACA CGC TTG GTP AND COG CTA CTC CCC GGC CTC TTG 2422 2316 ACC GAZ ATC CGC CTG CAT GCT GTA CGC ACA CGC TTG GTP AND CAC CTC CTC GGC CTC TTG 2423 2318 TIU CAR TGG GCA TGC CTG ATG CTT GAA CGT CAC GGA CGT AAC CTC CTG AGC CCC GCC AND CTT 72 2318 TIU CAR TGG GCA TGC CTG GAC AAC TTC TGG TCC CGT AAC CTC CTG AGC CCC GCC AND CTT 72 2318 T R R A C Q D R P R S R R C T A C T C T C T C T C T C T C T C T C		2	1.	•	A	A	K AST	L	: 151 C	. 360	AII	. GIB	AGE	- CEA	TAA	مت	1110	XCXC	YŒY	wa			:228.	
2296 ACTTANATOCIANCOCACTAGGCACHITHE ATO MAN TOT AME ANT GOD CTO ATO STO AND 2141 2314 ACC GTC CTC GAT GET GTA GOC ATA GOC THE OFF AME COD GTA CTC CCC GOC CTC TTG 2423 2414 CDC GAZ ATC CCC CTC GAT GET GTA GAZ ATA GOC THE OFF AME COD GTA CTC CCC GOC CTC TTG 2423 2414 CDC GAZ ATC CCC CTC ATG CCT GAA CTT GAC GOA GOA CAT AMC THAT OR GAX AND TCT CTC CTC 2433 2416 CDC GAZ ATC CCC CTC ATG CCT GAA CTT GAC GOA GOA GAT AMC THAT OR GAX AND TCT CTC CTC CTC 2433 2418 TIC CCC GAZ ATC CCC CAG CAX AAC TTC TOC TCC CCT AAC GTC CTC AGC CCG GCC AAAC TTT 77 2418 TIC CCC CAT CLU CCT GAT GAX ATA TAA AAC AAC AAA AAA AAA AAA AAA AAC GCA TA ATC CCC CCC ACC CAX ATT TAA TCATCGCCTCGTATA ATC TGT CCAA ATT GAA CTC ACC GCA ATC CAX TTG AAC CAX TTG AAA AAA AAA AAA AAA AAA AAC GTC AAC CCC ATC CAA CTC AAC CCC ACC AAC CCC ACC A								_	_	•	_		_	-	•								:	
2236 ACTTAMATICCTAMCGCAGTCAGCACTITITY ATO ANA TOT AND ANT GOOD CTO ATC STC ATC CTC GOL 2141 2346 ACC GTC ACC CTC GAT GCT GAT GCT GAT GAC GAT GCC GTC GTC ATC GAL ATC CTC CTC GAC CTC ATC GAT GCT GAT GCT GAT GCT GAT GCT GAT GCC GCC GAT ACC GCC GAC CAC ATC GAT GCC GAT GCC GAT GCC GCC GAT ACC GCC GCC GAT ACC GCC GCC GAT GCC GCC GAT GCC GCC GCC GCC GAT GCC GCC GCC GCC GCC GCC GCC GCC GCC GC	222	2	ACC T	ATG	I	7 3/7;	, W.	. 161	CAI		icy	CVA	CZĘX	TCAT	CONI	MGC	ITIA	AIU	تانت	:AG	:	ATCAC	2295	
2364 ACC CTC ACC CTG GAT GET GTA GGC AEA GGC THAT GGC THAT AND GGC GTA CTC CCC GGC CTC TTC 1423 2414 TO V T L D A V G C TG AAA GGT GAA GGT AAC GGT AAC GGT ACC CCC GGC CTC TTC 1423 33 R D I R E R R C TG AG CGT GAA GGT GAC GGA ACC GTA ACC GGG ACC AGG CGC AGC CGG AGC GGT AAC GTG TTC TGC TCC GGC AGC CGG AGC GGA AAC TTC TGC TCC GGC AGC AGC GGA AAC TTC TGC TCC GGC AGC AGC AGC AGC AGC AGC AGC AGC A			_		-		-	-	_	•	•							• • •	•	-			3	
2344 ACC CTC ACT CTC GAT GET GTA GGC ATA GGC TTG GTT AND GGC GTA CTC CCC GGC CTC TTG 1423 2444 CGC GAT ATC CGC CTC AAG CCT GAA CGT AAC GTA ACC GAA GGA AAG AGC TTC CGC GGC CTC TTG 1423 2444 CGC GAT ATC CGC CTC AAG CCT GAA CGT AAC GTA ACC GAA GGA AAC AGC CTC AGC GGA AAC TTC CGC GAT ACC GTA AGC CGC AGC AAC AGC CTC TTG 1423 2484 TIC CGC CAT CUC CTC ATC CCT GAA CGT ACC GGA AAC GTA ACC GTA AGC CGC ACC AAC TTC TTG TACA AAC TTC TGC TCC CGT AAC GTC CTC AGC CGC ACC AAC TTC TTG AAC CTC ACC CGC ACC ACC ACC ACC ACC ACC ACC AC																								
A COL CAT ARC COC CTG AAG CTG CAA CTG CAA CTG CAC CAC CAC CAC CAC CAC CTG CAC CTG CAC CAC CAC CAC CAC CAC CAC CAC CAC CA											_	_	-	_	_		_	-	-	:	_	6	-4	
ALUG COC CAT ARC COC CTG ATG COT GAA COT GAC COA COT AAC CAT AAC CAT ARC CAT ARC CAT COC CAT CAT CAT CAT CAT CAT CAT CAT CAT CA	236	4	ACC 7	erc v	ACC	<u>نت</u> و	GAT	े व्हा	GTA	, gac	AZA	COC	-	-	ASC	œ		ctc	ccc	990	CTC	πe	2423	
21846 THU USE TOO GCA TOC CAG CAE AND THE TOO TEEL COOP AND GTO COOP AND COO GET AND THE 1323 THE THE TOO TEEL COOP AND GTO CAG COO GET AND THE 172 THE TOO TEEL COOP AND GTO COOP AND THE 172 THE THE TOO TEEL COOP AND GTO COOP AND THE 172 THE TOO TEEL COOP AND GTO COOP AND GTO COOP AND THE 172 THE TOO TEEL COOP AND THE TOO TEEL C	_	_	•	•		_	_		-	-	-	•	_	•	_	-	•	-	-	-	L	:	35	
21846 THU USE TOO GCA TOC CAG CAE AND THE TOO TEEL COOP AND GTO COOP AND COO GET AND THE 1323 THE THE TOO TEEL COOP AND GTO CAG COO GET AND THE 172 THE TOO TEEL COOP AND GTO COOP AND THE 172 THE THE TOO TEEL COOP AND GTO COOP AND THE 172 THE TOO TEEL COOP AND GTO COOP AND GTO COOP AND THE 172 THE TOO TEEL COOP AND THE TOO TEEL C	242	•		ದಾ	YZC	œ	CIC	YIG	्दु	ĠN	्र द्	CAC	COA	ध्व	MC	CMC	(35	exc	LTG	fer	cre	crc	2433	
2344 ACT CCC CAT CLU UCT GIT GAC AAT TAA TCATCGGCTCGTATA ATG TOT CCA ATT GIG AGC GGA 74AAC 773 T P 8 P P V U U T N S G T V S G T S S S S S S S S S S S S S S S S S	•	•	_	•	•	^	-			Ł	×	b	G	II.		ı		n	M	C	v	-	62	
2344 ACT CCC CAT CLU UCT GIT GAC AAT TAA TCATCGGCTCGTATA ATG TOT CCA ATT GIG AGC GGA 74AAC 773 T P 8 P P V U U T N S G T V S G T S S S S S S S S S S S S S S S S S	218	•	TTC	æ.	766	CY	700	CNG	CVE	w	110	706	700	cat	AAC	CTC	CTG	AGC	œ	ann.	MAC:	~;;	2542	
2607 TAA CAATTECKONGNOAMACAGGREE AND CAT ATO MAA AMA CAT AND AMA AMA CAT CAM AND AND CAG GTA ATG 26'0 B B B B C CT CTO CTO MAA GAA TOC CAG CTG ATC CAA CTG AMC 2710 B B B B C C C C C C C C C C C C C C C	•	•		•	-	~	•	4	U		•		3	Œ	ı	V	L	3	2	A		r.	72	
2607 TAA CAATTECKONGNOAMACAGGREE AND CAT ATO MAA AMA CAT AND AMA AMA CAT CAM AND AND CAG GTA ATG 26'0 B B B B C CT CTO CTO MAA GAA TOC CAG CTG ATC CAA CTG AMC 2710 B B B B C C C C C C C C C C C C C C C	234	4 /	MCT	œc	CAI	تانت	Œ	GII	c)C	MI	TAA	100	tesa	CTC5	AZA	ATG	197	œ	ATT	GTG	L EC	æı	2606	
2671 AMA CHAPTTENCAGRADAMACAGGRIPE ATG CAT ATG MAA AMA ANC ATC AMA AMA AMC CAG GTA ATG 2670 2671 AMA CTC CCT CCC AMA TCT AMA CTC CTC AMA GAA TAC AMA AMA CCC CAG CTG ATC CAA CTG AMC 2711 W L C P W B R A C L R R Y R A W L L R R R P Q V N H 2731 ATC CAA CMA TTT GAA GCA CCT ATC CCT CTG ATC CTG BET GAT BCT TAC ATC CCT TUT CCT 2731 ATC CAA CMA CMA TTT GAA GCA CCT ATC CCT CTG ATC CTG BET GAT BCT TAC ATC CCT TUT CCT 2731 CAT CAA CMA CCT AAA ACC THC TCT ATG CAC TTC CAC CCC CCC CCC CCC CCC AAA AMA AMA GCA CAT ATC GAC CAC 2830 2731 CAT CAT CTG CTG TAC CAT CAG TAG CTG CTG CTG CCC CCC CCC CCC CAC AAA AMA GAA CAT CTT AMC 273 V C L L T D Q W V L R B P R R R R R V B 2731 CAC CTG CCT AMC CTG CTG TAC CTG TAC CTG CGC CCC CCC CCC CCC CAC AAA AMA GAA CAT CTC AMC 273 V C L L T D Q W V L R B P R R R R R V B 2731 CAC CTG CCT AMC CTG CTG TAC CTG TAC CTG CGC CCC CCG ACT TTC AMA CAC CAA CCT TCC AMC 273 V C L L T D Q W V L R B R R R R R R V B 2751 AMA CTG CCT AMC CTG CTG TAC ACC TCG CGC CCC CCG ACT TTC AAA CAC CTG CAC CAA CAC TCC 2751 AMA CTG CCT AMC CTG TTC ATC CTG TAC AMC AMA AMA ACC ATC CCC CCC CCC AAC AMA AMA CCC TCC CTG CTC CAA 2751 AMA CTG CCT AMC CTG TAC ATC CTG CTG AMC AMA AMA ACC ATC CCC AAC AMA CAC CTG TTC CAA 2751 AMA CTG CCT AMC CTG TCC ATC CTG CCA CCC CAG ACT TTC ATC CTG CTG CAA 2751 AMA CTG CCT AMC CTG TCC ATC CTG CCA CCC CAG ACT TCT TCC ACC TTC CAA 2751 AMA CTG CCT AMC CTG ATC CCC CCG ATC TCC TCC CAG CCC CAG ACC TCT TCC ACC CTG TCC TCC CAC CCC 2751 AMA CTG CCC CCG ATC TCT CTG CCC ATC CTG AMC ACC CAG TCT TTC ACT TTC CAA 2751 AMA CTG CCC CCG ATC TCT CTG CCC CAG AMC AMA ACC CCG CCC CAG TCC TTC CAC CTG TCC TTC CAC CTG TCC TCC CAC CCC CAG AMC AMA TCC CCG CCC CAG AMC AMA TCC CCG CCC CAG AMC AMA TCC CCG CCC CAG AMC AMA TCC CCC CAG AMC ACT TCC TCC CAC CCC CAG AMC AMA TCC CCC CAG AMC ACT TCC TCC CAG CCC CAG AMC AMA TCC CCC CAG AMC AMC TCC CCCC CCCC CCC CCC CCC CCC CCC CCC	7.	3		•		7	•	▼	U		•				-	×	¢							
2731 AMC CTC CCT CCC AMC TCT AMA CTC CTC AMA GAA TMC AMA TCC CMG CTG ATC CAA CTG AMC 27310 13 U L G P H B H L L H B Y H A U L T H B A U L T H B A U L T H B A U L T H B A U L T H B A U L T H B A U L T H B A U L T H B A U L T H B A U L T H B A U L T H B A U L T H B A U L T H B A U L T H B A U L T H B A U L T H B A U L T H B A E E E E E E E E E E E E E E E E E E	2607 TAA CAAPTTCACACACACACACACACACACACACACACACACA											2410	-											
2731 ATC GAA CAG THE GAA GCA GCT ATC GCT CTC ATC CTC CTC GAT GAT GCT TAC GCT T		_	•									n	£		1	X					×		•	ı
2731 ATC GAA CAG THE GAA GCA GCT ATC GCT CTC ATC CTC CTC GAT GAT GCT TAC GCT T	267	1 /	uc	cze	and i	cce	AAC	101	***	٠	cte	***	EAA	TAC	***	TCC	CMG	CTG	ATC	CAA	CZS	м	2735	i
2791 GAT GAA GGT AAA ACC THC TGT ATG CAG TTC CAG TCC ANA AAC AAA GGA TRC ATG GAG CAC 2830 53 D E G E T Y C H O F E W R H R A Y H O R 72 2831 GTA TGT CTG CTG TAC CAG TGG TGG TGG CCC CAG AAA AAA GAA CAT GTT AAC 2910 73 V C L L Y D Q W V I. R B P H R H R A R CAC TGG CGT AAC CTG GTA ACC TGG GGC GCC CAG ACT TTC AAA CAC CAT GCT AAC CTG GTA AAC CTG AAC CTG AAC CTG AAC CTG AAC CTG AAC CTG CAC GTA AAC CTG CTG AAC CTG CTG AAC CTG CTG CTG CAC CTG CTG CTG CAC CTG CTG CTG CAC CTG CTG CTG CTG CTG CTG CTG CTG CTG CT	'	۱ ۴	J	L	c	P		•	K	L	L	X		T			ų		r		L			
2791 GAT GAA GGT AAA ACC THC TGT ATG CAG TTC CAG TCC ANA AAC AAA GGA TRC ATG GAG CAC 2830 53 D E G E T Y C H O F E W R H R A Y H O R 72 2831 GTA TGT CTG CTG TAC CAG TGG TGG TGG CCC CAG AAA AAA GAA CAT GTT AAC 2910 73 V C L L Y D Q W V I. R B P H R H R A R CAC TGG CGT AAC CTG GTA ACC TGG GGC GCC CAG ACT TTC AAA CAC CAT GCT AAC CTG GTA AAC CTG AAC CTG AAC CTG AAC CTG AAC CTG AAC CTG CAC GTA AAC CTG CTG AAC CTG CTG AAC CTG CTG CTG CAC CTG CTG CTG CAC CTG CTG CTG CAC CTG CTG CTG CTG CTG CTG CTG CTG CTG CT	273	1 /	NTC	an.	cac.	777	GAA	ŒA	SET.	ATC	cc	CTC	ATC	C70	-	GAT	ec t	TAC	ATC	267	~~	O (20)	2780	-
2851 GTA TOT CTG CTG TAC GAT CAG TIM GTA CTG TCC CCC GCC CAC AAA AAA GAA CGT GTT AAC 2910 73 V C L L I D Q H V I. R B P H R H B R V H 92 2851 CAC CTG CGT AAC CTG GTA ATC ACC TCG GCC CCC CAG ACT TTC AAA CAA GAA CGT GTT AAC 2910 93 H L G H L V I T H G A Q I F H R Q A F H 112 2971 AAA CTG GCT AAC CTG TTC ATC GTT AAC AAC AAA AAA AAA ACC ATC CCG AAC AAC CTG GTT CAA 1020 113 H L A H L F I V H H R R I I F H H I. V F 133 1031 AMC TMU LTG ACC CCG ATG TCT CTG GCA 2AC TOG TTC ATG GAT GGT GGT AAA TGG GAC 3690 133 H I L I F H R I L A Y H F H D D G G R H D 159 3091 TMC AMC AAA AAC AAA TCG ATC GTA CTG AAC ACC CAG TCT TTC ACT TTC GAA 3150 153 I V H R II I V R G L R H R V Q L H C Y V R 192 23181 CAA GTA GAA AAC AAA CGG ATC TAC GAT GGT AAL AAC TCG AAC TCT TAC GTA AAA 3210 173 R V R Y L V R G L R H R V Q L H C Y V R 192 2311 ATC AAC AAA AAC AAA CGG ATC AAC ACC CAG ACT TAL UTG ATC TTC TAC AAC 3270 189 I H R H E P I I Y I D I H I I I I I I I I I I I I I I I I] a:	3 1	7									L		L	4	_	_ ~			Ħ.	_			1
2851 GTA TOT CTG CTG TAC GAT CAG TIM GTA CTG TCC CCC GCC CAC AAA AAA GAA CGT GTT AAC 2910 73 V C L L I D Q H V I. R B P H R H B R V H 92 2851 CAC CTG CGT AAC CTG GTA ATC ACC TCG GCC CCC CAG ACT TTC AAA CAA GAA CGT GTT AAC 2910 93 H L G H L V I T H G A Q I F H R Q A F H 112 2971 AAA CTG GCT AAC CTG TTC ATC GTT AAC AAC AAA AAA AAA ACC ATC CCG AAC AAC CTG GTT CAA 1020 113 H L A H L F I V H H R R I I F H H I. V F 133 1031 AMC TMU LTG ACC CCG ATG TCT CTG GCA 2AC TOG TTC ATG GAT GGT GGT AAA TGG GAC 3690 133 H I L I F H R I L A Y H F H D D G G R H D 159 3091 TMC AMC AAA AAC AAA TCG ATC GTA CTG AAC ACC CAG TCT TTC ACT TTC GAA 3150 153 I V H R II I V R G L R H R V Q L H C Y V R 192 23181 CAA GTA GAA AAC AAA CGG ATC TAC GAT GGT AAL AAC TCG AAC TCT TAC GTA AAA 3210 173 R V R Y L V R G L R H R V Q L H C Y V R 192 2311 ATC AAC AAA AAC AAA CGG ATC AAC ACC CAG ACT TAL UTG ATC TTC TAC AAC 3270 189 I H R H E P I I Y I D I H I I I I I I I I I I I I I I I I	279	. (31	en.	ŒĨ	**	ACC	TAC:	-6-	ATE	243	TTC	cac	700	***	e	**	~~	-		- 2-		100	1
2911 CAC CTG GGT AMC CTG GGA ATC ACC TGG GGC GCC CMG ACT TTC AMA CAC GGA GCT TTC AMC 3070 93 E L G H L V I I H G A Q I F E R Q A P H 112 2971 AMA CTG GGT AMC CTG TGC ATC GTT AMC AMC AMA AMA ACC ATC CGG AMC AMC CTG GTT CMA 1030 113 E L A H L F I V H H R I I P H H I. V F 132 2031 AMC TRU CTG ACC CGG ATG TCT CTG GGA 2Mc TGG TTC ATG GAI GGT GGT AMA TGG GAT 3600 133 H L I F H H S L A Y H F H D D G G R H D 157 3091 TMC AMG AMA AMC TGT AMC AMC AMA TGG ATC GTA CTG AMC ACC CMG TCT TTC ACT TTC CMA 3150 153 H F H H S L A Y H F H D D G G F H D 157 2181 CAA GGA GGA GAA TAC CTG GGT AMA GGT CTG CGG AMC ACC CMG TCT TTC ACT TTC CMA 3150 173 H V H R G L R H M M M M M M M M M M M M M M M M M M	5:	1)	£	G	X	7	Y	C		0	r		•										1
2911 CAC CTG GGT AMC CTG GGA ATC ACC TGG GGC GCC CMG ACT TTC AMA CAC GGA GCT TTC AMC 3070 93 E L G H L V I I H G A Q I F E R Q A P H 112 2971 AMA CTG GGT AMC CTG TGC ATC GTT AMC AMC AMA AMA ACC ATC CGG AMC AMC CTG GTT CMA 1030 113 E L A H L F I V H H R I I P H H I. V F 132 2031 AMC TRU CTG ACC CGG ATG TCT CTG GGA 2Mc TGG TTC ATG GAI GGT GGT AMA TGG GAT 3600 133 H L I F H H S L A Y H F H D D G G R H D 157 3091 TMC AMG AMA AMC TGT AMC AMC AMA TGG ATC GTA CTG AMC ACC CMG TCT TTC ACT TTC CMA 3150 153 H F H H S L A Y H F H D D G G F H D 157 2181 CAA GGA GGA GAA TAC CTG GGT AMA GGT CTG CGG AMC ACC CMG TCT TTC ACT TTC CMA 3150 173 H V H R G L R H M M M M M M M M M M M M M M M M M M	285	1 6	77 A	197	CTG	CTG	114	G T	cas	736	CT3	~**	***		~~	-								1
2911 CMC CTG CCT AAC CTG GTA ATC ACC TGG CGC GCC CAG ACT TTC AAA CAC CTA CCT TTC AAC 1070 93 E L G H L V I T H G A Q P P E R Q A P H 112 2971 AAA CTG GCT AAC CTG TTC ATC GTT AAC AAC AAA AAA ACC ATC CCG AAC AAC CTG GTT CAA 1010 113 E L A H L P I V H H E R T I P H H I. V P 1312 3051 AAC TAL CTG ACC CCG ATG TCT CTG GCA 2AC TCC TTC ATG CAT CAT CTG TTC AAC TTC CAA 1020 133 H T L P H R S L A Y H P H D D G G Z H D 1557 3091 TMC AMC AAA AAC TLT AMC AAC AAA TCG ATC GTA CTG AAC ACC CAG TCT TTC ACT TTC CAA 1150 133 F F H R H S Z H R S I V L H P Q L H C Y V R 1172 3181 CAA GTA GTA GAA TAC CTG GTT AAG GTT CTG CGC AAC TCC TAC CTG TAC CTA AAA 1210 173 E V B Y L V E G L R H E P Q L H C Y V R 192 2211 ATC AAC AAA AAA CGG TAC CTG ATC TAC ATC GAA TCT ATG TCT TAL CTG ATC TTC TAC AAC 1210 183 I H R H E P Y L I P Q H H Y R L P H T I S S 232 3271 CTG ASC AAA CCG TAC CTC ASC CCC CAG ASG ANG TAC TAC GCC AAC ACT ATC TCL TCL TCL TCL TCL TCL TCL TCL TCL T	7	3 1	,	c	L	L	1	5	•		v	1.										ME I		Ì
2971 AAA CTG GCI AAC CTG TIC ATC GTT AAC AAC AAA AAA ACC ATC CCG AAC AAC CTG GTT GAA 1010 113 E L A B L P I V B B E R T I P B B II. V P 132 3031 AMG TAU CTG ACC CCG ATG TCT CTG GCA 2AC TOG TTC ATG GAT GAT GGT AAA TGG GAP 3090 133 B T L T P B S L A Y B P B D D G G T AAA TGG GAP 3090 133 B T L T P B S L A Y B P B D D G G T AAA TGG GAP 3090 133 B T L T P B S L A Y B P B D D G G T AAA TGG GAP 3150 3091 TAG AMG AAA AAC TUT AAC AAA AAA TCG ATC GTA CTG AAC ACC CMG TCT TTC ACC TTC CAA 3150 133 B Y B B S T B B S T B B S I V L B T G S P T P B 172 3181 CAA GTA GAA TAC GTG GTT AAG GGT CTG CAA CTG AAC TGT TAC GTA AAA 3210 173 B V B Y L V R G L R B E P G L R B E P G L F T B 312 3271 CTG AGC AAA AAA CGG TAC CTC ATC CCC CAG ANG ATG TAC TCT TAU UTG ATC TCC TAC TAC TCC TAC TCC TAC TCC TAC TCC TAC TCC TC	2911			c 7 a	<u>مت</u>	116	- 10	47 2	1 ~	100	***	cee	معي					***						
2971 AAA CTG GCT AAC CTG TIC ATC GTT AAC AAC AAA AAA ACC ATC CCG AAC AAC CTG GTT GAA 1030 113 E L A B L F L V E H E R T L F E H H 1. V F 132 1330 1331 AAC TAM CTG ACC CCG ACC ACC ACC ACC ACC ACC ACC AC	93	ì	Ī	:	<u>.</u>	1	ī	Y	Ĩ	ī	*	~						R			_			
113 A B B B B C COR ARE TOT CTG GCA 2AC TOG TTC ATG GAT GAT GGT AAA TGG GAT 3G00 133 B T L T P B S L A Y B P H D D G G T TTC ACT TTC GAA 3150 133 B T L T P B S L A Y B P H D D G G T TTC ACT TTC GAA 3150 133 B T L T P B S L A Y B P H D D G G T TTC ACT TTC GAA 3150 133 B T L T P B B S T B B S T B B S T P P B 172 1151 GAA GEA GAA TAG GTG GTT AAG GGT CTG GGT AAL AAA TTC GAA CTG AAC TGT TAC GTA AAA 3210 173 B V B Y L V R G L R B E P G L R B E P G L B C Y V E 192 1251 ATC AAC AAA AAC AAA GGG AFC ATC TAC ATC GAT TGT ATG TGT LAL UTG ATC TTC TAC AAC 3270 183 C A R B E P T L T T T T T T T T T T T T T T T T T	207		LA.A	c†a		110					115			-	-	•	_		-		- .			1
JOST AME THE LITE ACC CCE ATE TOT CTG GCA ZAC TOG TTC ATG GAT GAT GGT GAT TAG GAT 1330 B T L T P H S L A Y B P H D D G G T AAA TGG GAT 1400 1537 B T B T B T B T B T B T B T B T B T B	111	i	į.	ī	Ÿ.	T	L	ř	Ī	A													- · ·	
3091 THE AME AMA AME TUT ALC AME AMA TOG ATC GTA CTC AMC ACC CAG TOT TIC ACT TTC CAA 3150 153 Y W W W S Y W W W S I W W W T Q S F T P 8 172 2181 CAA STA CAA TAC GTT AMA OST CTS CAI AMU AMA ITC CAA CTG AMC TOT TAC GTA AMA 3210 173 W W W Y W W W W W W W W W W W W W W W	1			924 •								• • • •		_	•	-	•	-	_		•	-		
3091 The AME AMA AME TOT ME AME AME AMA TOG ATE GTA CTG AME ACC CAG TOT THE ACT TTC CAA 3150 123 Y . W H H S Y H B S Y H W S S Y H W H S Y H S S S S S S S S S S S S S S S S	131	í		T.	L	1	7	N.A.	161 5	L		Y										_		
2181 CAA SEA SAA TAG CTG SET AAS SET CTS CGI ANU AAA ITC CAA CTG AAC TGI IAC GIA AAA 3210 173 8 V 8 Y L V R G L R 8 E V G L W C Y V E 192 2211 ATC AAC AAA AAC AAA COS AFC ATC IAC AEC SAI TGI ATG TGI IAL UTG ATC ITC IAC AAC 3210 183 I W R W W R R R R R R R R R R R R R R R	,		-						•	_		_	•	-	••	-	-	•	-	-	-	-	10	
3181 CAA SEA SAA TAG CTG SIT AAS SET CTS CGI AAU AAA ITC CAA CTG AAC TGT IAC GIA AAA 3213 173 8 V 8 Y L V R G L R 8 E Y Q L W C Y V R 193 2213 ATC AAC AAA AAC AAA COS AIC AIC IAC SAI TGT AIG TGT IAU UTG ATC ITC IAC AAC 3270 183 I 8 R 8 E P I I Y I D S N 8 Y L I Y I N 212 2213 1. T E P Y L I P R 212 2213 1. T E P Y L I P R N N N T R L P N I I S S 212 2213 1. T E P Y L I P R N N N T R L P N I I S S 212 2213 1. T E P Y L I P R N N N T R L P N I I S S 212 2213 1. T E P Y L I P R N N N T R L P N I I S S 212 233 1. T E P N T L I P R N N N T R L P N T I S S 212 233 1. T E P N T L I P R N N N T R L P N T I S S 212 233 1. T E P N T L I P R N N N T R L P N T I S S 212 233 1. T E P N T L I P R N T R L P N T L I S S 212 233 1. T E P N T L I P R N T R L P N T L I S S 212 233 1. T E P N T L I P N T R L P N T L I S S 212 233 1. T E P N T L I P N T R L P N T L I S S 212 233 1. T E P N T L I P N T R L P N T L I S S 212 233 1. T E P N T L I P N T R L P N T L I S S 212 233 1. T E P N T L I P N T R L P N T L I S S 212 233 1. T E P N T L I P N T R L P N T L I S S 212 233 1. T E P N T L I P	יננ) 1		7	~~		17-17 8	7													_	_	1 1 1 1	
2311 ATC AMC MAA AMC MAA COS ATC ATC TMC MAC GAS TOT ATG TOT TML LTG ATC TTC TMC AMC 3270 183 L H R H E P I I Y I D S H S Y L I Y T H 312 3271 CTG ASC MAA COS TMC CTC ASC CCC CMS ASS ATS TMC AMA CTG CCC AMC ACT ATL TLL TXL 3330 313 L T E P Y L I P Q H H T R L P H T I S S 232 3331 GMA ACT TTC CTG AMA TMA TAMSTCCACCTCCAGCCCACCTCCCCAGCCCACCTCCCCAGCCCACCTCCCCAGCCCACCTCCCCAGCCCACCTCCCCAGCCCACCTCCCCAGCCCACCTCCCCAGCCCACCTCCCCAGCCCACCTCCCCAGCCCCACCTCCCCAGCCCCACCTCCCCAGCCCCACCTCCCCAGCCCCACCTCCCCAGCCCCACCTCCCCAGCCCCACCTCCCCAGCCCCACCTCCCCCAGCCCCACCTCCCCAGCCCCACCTCCCCAGCCCCACCTCCCCAGCCCCACCTCCCCCAGCCCCACCTCCCCAGCCCCACCTCCCCCAGCCCCACCTCCCCACCTCCCCAGCCCCACCTCCCCAGCCCCACCTCCCCAGCCCCACCTCCCCACCTCCCCACCTCCCCACCTCCCCACCTCCCCACCTCCCCACCTCCCCACCTCCCCACCTCCCCACCTCCCCACCTCCCCACCTCCACCTCCCACCTCACCAC	1			_	_	_	-	-	_		_	•	•	_	_		•	_	•	•	•	•	4/6	
1931 ATC ANC MAA ANC MAA COS ATC ATC TAC AND ATC TOT ATG TOT TALL UTG ATC TTC TAC ANC 1270 193 I H R H E P I I T T I O S H S T L Y T H 212 3271 CTC ATC ATC MAA COC TAC CTC ATC CCC CAS AND ATC TAC ANA CTC CCC ANC ACT ATC TUL TOL TOL 1330 213 I. T E P Y L I P Q H H T R L P H T I S S 232 1331 GNA ACT TTC CTG ANA TAN TANGETCENCOTCANCOTCANCESCOSTETIACANCESTOSTENCE 3404	175	. (Y V	5 \	TAC																		-
3271 CTG ARE AAA CCG THE CTC ARE CCC CAS ARE ARE THE AAA CTG CCC AAC ACT ATC TCU TX: 3330 213 T. T E P Y L I P Q N N T X L P N T I B S 232 2331 CMA ACT TTC CTG AAA TAA TAA TAACTCCACCTCCACCCACCCACCTCCACCTCCCACCTCCCACCTCCCACCTCCCACCTCCCACCTCCCACCTCCCACCTCCACCA	1				-	•	_	-	-	_		•	-	_	-	_	_	_	-	_	•	_	. 76	
3271 CTG ASC AMA CCG THE CTC ASC CCC CAG AGG ATG THE AMA CTG CCG AME ACT ATG TOU	2323		75																ATC			_		1
3331 GM ACT TTC CTG ANA TAN TANGETCENCETCONGCONCATCONCETCONCETCATTENCHACTORS AND TANGETCENCETCON				_		_	_	-	_	_		_	-	_		_	•	_	•	•	-	-	414	-
3331 CAA ACT TTC CTG AAA TAA TAACTCCACCTCCACCCCACC	3271		76																					
					_	•	-	_									_							-
	3333	•	A.	KT	TTC	CTG	144	738					200	CCT	اعتما	cta)CC(E	rce#1	ITZM	:AACI	77 CE	PEACT	3404	
	133	, [1	<u> </u>	L	1	•															Z 34	,

I-SceI coding sequence of pSCM525 - Note the two amino acid. N-terminal extension as compared to genuine version of the gene.

YARIATIONS AROUND THE 1-SCR | SEQUENCE

		-2 H_	-1 -	1	r	<u> </u>	_1_	5		ч	_1_		10	4	L	2	þ		•
K	ţ		20 K		Y				L				30 N		Ē	•	•	į	4
G	I	£	40	ı	<u>_</u>	ā	2	A	I	1	ર	S	∞	c	ξ	3	(•	1
C	Ħ	٥	60 F	Ε	W	ĸ	N	ĸ	A	Y	м	0	70 H	۷.	, c	L	-	v	:
Q	٧	¥	80	3	P	٠ ٩	н	ĸ	K	٤	R	¥	90 N	Ħ	L	a	N	_	Ą
1	T	٧	100 G	A	٥	T	F	ĸ	н	a	A	F	110 N	K	٠ ر	A	N	-	:
1	٧	N	120 N	K	K	I	ı	P	N	N	L	٧	130 E	N	Y	L	ī	2	•
3	L	A	140 Y	w	Γ	n	2	<u>0</u>	2	2	*	٧	150	Y	N	K	N	s	ï
N	K	\$	1 60 1	٧	L	N	T .	Q	\$	۴	t	F	170 E	Ε	٧	Ε	Y	-	٧
ĸ	e	L	180 R	N	K	F	٥	L	N	c	Y	٧	1 90 K	I	N	K	N	<	,
1	1	Y	200)	8	Ħ	s	Y	L	1	F	Y	210 N	L	t	K	P	4	•
I	P	Q	220 M		Y	K	Ĺ	P	N	T	t	3	230	E	T	F	Ĺ	K	

Positions than ean he changed without affecting mayone activity (demonstrated)

positions -1 and -2 are not natural. The two amino acids are added due to cloning strategies

positions 1 to 10: can be deleted

position 36: G is tolerated

position 40: M or Y are tolerated

position 41: 8 or M are telerated

position 43: A is tolerated

position 44: Yer Mare tolerated

position 91: A is tolerated

positions 133 and 150: Lare tolerated position 233: A and Sare tolerated

Changes that affect engyme activity (demonstrated)

position 19: L to &

position 38: 1 to 5 or M

position 39: G to D or R

parilion 40: L to Q

position 42: L to B

position 44: Die E. Cor H.

pesition 45: A to E or D

profition 46: Y to D

position 47: Its B or N

position 50: LwS position 144:DwS

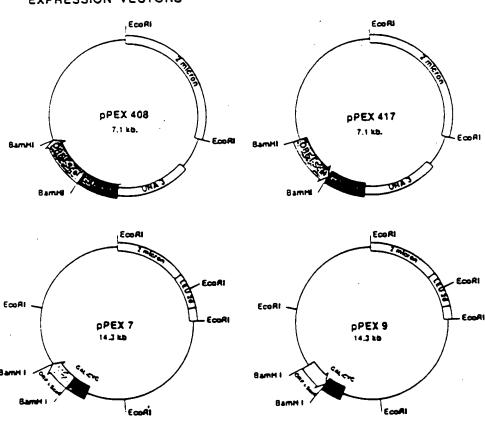
postice 145:D to E

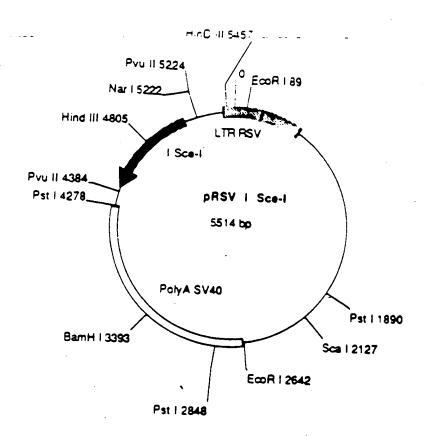
position 146 Q to E

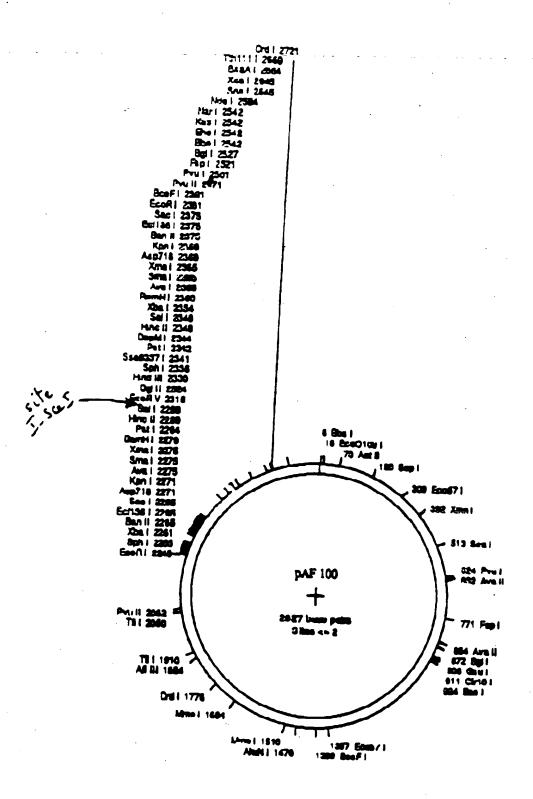
Group I Intron Encoded Endonucleases and Related Endonucleases

				<u></u>
_		indomiclease	Recognition Sequence SITE	▼Intron site
	(Sectheron	I-6cs I Tyces mitochondria)	CGCTAGGGATAACAGGGTA	ATATATE
	(Seccheron	I-Sca IV nyces mitochondria)	AAGAGTACTAA LUAGA EE	
E	(Seccharoz	l-Scall Nyces mitochondria)	GAAACCAGTAGAGTAGAAGTA	*
the cutt	(Chlanydo	I-Cau I mouse chlorophat)	TARCOGICCA A GOTAGESA ATIGCCA GATTCCA TCGC	
139	(Ph y	1-Ppo (Serum nucleus)	T G A C T C T T A A G C T A G C C A G C C A G C A G A A F F C C A T C G G G	A & C C C C
Pamily	(Seccharon	I-See III yess snitochondria)	C T C C A A A U C A T T A A A A A A A A A A A A A A A A	ATTACE
patide	(Chiamydo	I-Cre I monas chloropiast)	GGGTTCAAAACGTCGTGAGA CCCAAGTTTTGCAGCAC+C+	7 7 7 9 A 9
Two Dodecapatide Pamily for 4 hp cuttern)	Caccharon	See I (RF3) yeas miliochondria) is resid	GATOCTOTAGOCATAGOCT CTACGACATOCOTATAGOCT	1 4 1 7 8 9
2	(Saccherr (Nos	HO Impose nucleus) Incomé	CTTTCCGCAACAGTATAATT	7 7 A 7 A A A A A A A A A A A A A A A A
		I-Cam I nonas mitochondria) endemodesse)	ACCAIGGGG TCAAATUTCIT TGGTACCCCAGITTACAGAA	TCTEGG AGACCC
		I- Pan I mitochondria)	GIGCCTGARTGATATHAT + CACCCACTCACTATAAAAA	FOGNAN
Ma	Becomophi	ge TO	-	
Fem	l Tev l	GAAC GC CA G T C	TAGATGITTCITGGGCTACC	OTITAAT CAAATTA
Cuchura	l-Tev II	CAAG		6 1 1 1 7 7 6 1 1 7 1 1
Other Servetural Fa	l-Tev III	GCTATTCGTT	tiltare three transcare the table to table to table to the table to table	6 2 2 2 7 7 1 1
7				

EXPRESSION VECTORS







```
par 100 - RESTRICTION MAP
```

```
Sau3A I
                                       HOO I
                                   Sci I
                                   Mop I
                                   Epa ::
                                   V KEC
                                   3328 :
                                  Sha I
                                      Bety I
                        Halk I
                                  Batk I
                        Bap1286 I Bon I
                                  Bon I
                             3an I
                                                 Rine II
                                               I Ega I
                                  AVE 1
                         ALU I
                              Asp718 ALW I
              Nap I Rea I
                                     11 1 1
                                                 11 1
                              11 11
                    11 11
CCAMCTTCAATTCCCATCCTCTAGACCTCCCTACCCGGGATCCTCCAGTCAACCC
COTTOCACOTTA ACCOTTA COSACA TOTOCA COCATOCOCOTA COMO COCATOCO A TOCOTA TOS COULA CIA A UTOTA
                   •11 11 •11 L1 L1 L
                                                 11 1
   1 110
              11
                                                                               2318
                                      2279
                              2271
                                                 2286
              2255
   2244
                                                   2292
                                   2275
     2247
                                     2279
        2249
                                   2275
              2255
                                            2284
                                                 2289
                                   2276
                         2245
              2255
                                                  2290
                         2265
2460
                                   2275
               2256
                                   2273
                         2265
                                       1280
                                2272
                                       2279
                               2271
                               2271
                                       2279
                                2272
                                        7790
                                   2275
                                   2275
                                   2275
                                    2276
                                     2276
                                    2276
                                     2274
                                        2280
                                         2280
                                        2280
```

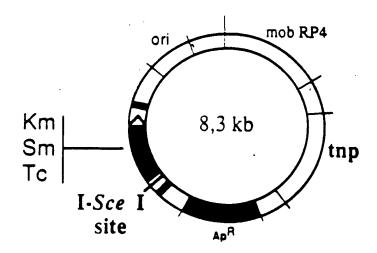
paf 100 - RESTRICTION MAP

```
Ser! I
                                               Yci:
                                               S.F.
                                               374
                                               234 V
                                               SALK I
                                               2C1 :
                                              SCEP I
                                              BACK T
                                              BaaJ ?
                                              Ben :
                                          SAUJA I
                                                        Sac 1
                                         Mbo I
                                                        EgiA I
          Alu I
                                         Don II
                                                        reille :
                     ste :
                               Ple I
                                         Opa I
                                                  Rea I
       Dde I
                     Pat I
                               Bin: :
                                        NLA IV
                                                  Capé I
    1 AUDSZ
                     See 337 1
                                                 AIT, LA
                                        I YseE
    EDO I
                NIA III
                                        I have
                                                 Kpa
    DDE IT
               Spn I
                            Sal I
                                        Alw I
                                                 Bas I Alu I
    Opa I
                                                                       H ag
                           mine :I
                                      MAL I AVE I
                                                       9sp1286 [
   BULT I
                                                                      SOL ::
               #ap7524 I
                           ACC I SEE I
                                           RESU 1
                                                       San II
                                                                      244
                                  ALW 1 BOOK
                                                 ASP/LE
                                                           Tag I
                                                                    Bar I
   11 1 11
               11
                                     an in in
A TONONTO INNOCTITO CATOCCTO CAROCT CTAGAGGA TO COORGET ACLUACUTURA VICALIGACIO COTT
THE TET HEATTCOANCOTACOGACOTCCAGC TO A GATET CETAGOGCCCA TODCT COASCITANG TO ACCUMENTAL AND
               11 -11 1
                           11-1
                                     3 (1 ) 111 (1)
                                                       11 :-1
   2324
                                                                   1 11
               2336
                       2311
                                 2351
                                        2261
                                                 2360
                                                           2379
                                                                   2387
   2324
               2336
                           2246
                                  2356
                                            2364
   2325
                                                                     2369
               2336
                           2348
                                     2350
                                            3.3 CB
                                                       2375
                                                                     2369
    2329
               2336
                           2348
                                       2360
                                                 2369
                                                        2376
                                                                      2330
   2325
               2337
                            2344
                                       2360
                                                 2360
                                                             2301
                                                                       2391
   2325
                    2341
                                       2760
                                                 2368
      2328
                     2347
                              2351
                                       2760
                                                  2370
        2330
                     2342
                              2351
                                        2361
         2331
                                        2361
                                                       2375
                                        2361
                                                       2375
                                        2361
                                                       2375
                                            2343
                                            2365
                                            2365
                                            2365
                                            2365
                                            2365
                                            236
                                            2365
                                             2366
                                             2366
                                             2366
                                             2366
                                             2346
                                             2366
                                             2366
                            Mae III
                      Serf I
                      ROTER T
                                                                       P~ II
                                                                      Nama II
                     terencos un igesticanting estance testeturg esselung este electrone
   2405
                     2423
                                       2440
                                                  2451
                                                                      2471
        2416
                      2424
                                                  2451
                                                                      2471
                      2424
                                                        2457
                                                                       2472
                      2424
                      2424
                      2424
                            2430
```

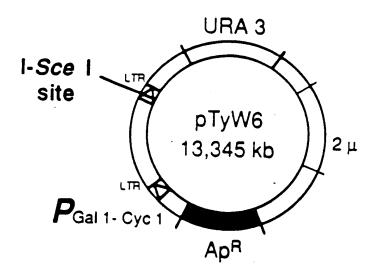
Names: $pTSm \omega$

pTKm ω

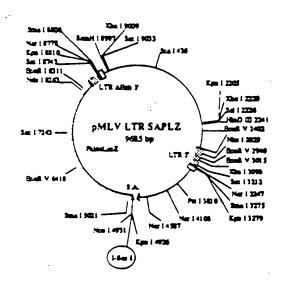
ρΤΤς ω

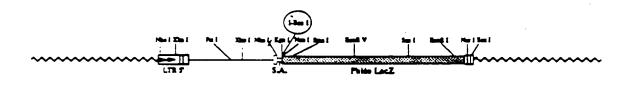


Construction: pGP 704 from De Lorenzo, with transposase gene and insertion of the linker [I-Scel] in Notl unique site

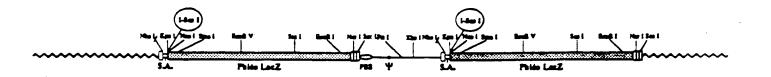


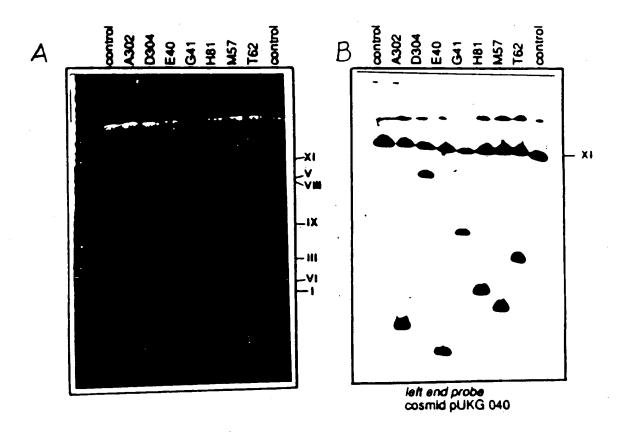
Construction: pD 123, from J. D. Boeke with insertion of a linker [I-Scel - Notl] in BamHI



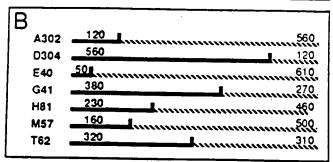


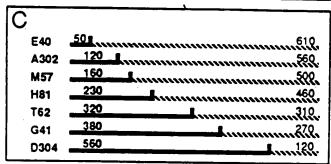


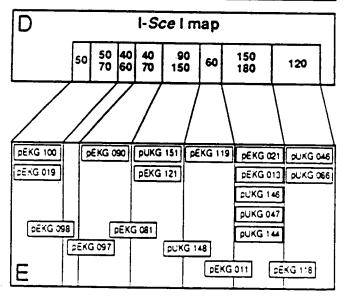


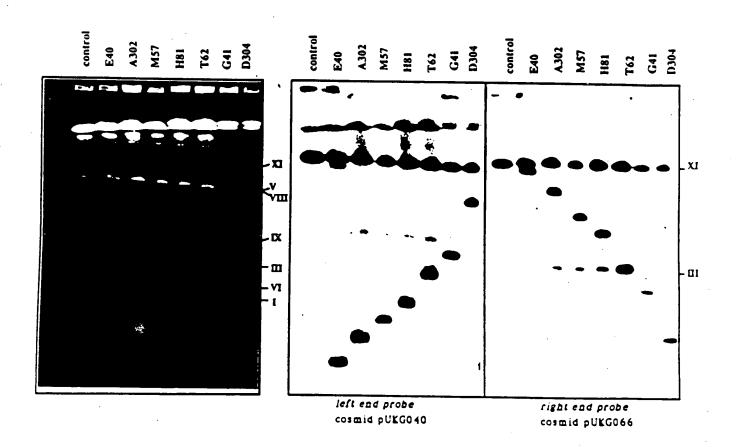


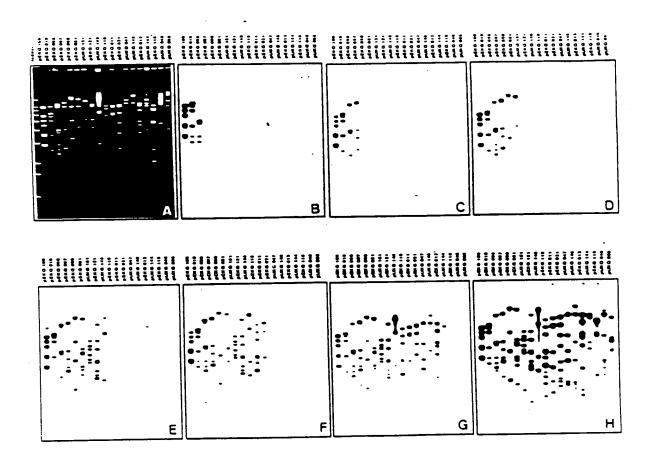
A	
A302	Minimum minimum Milim
D304	
E40	
G41	270
H81	
M57	500
T62	

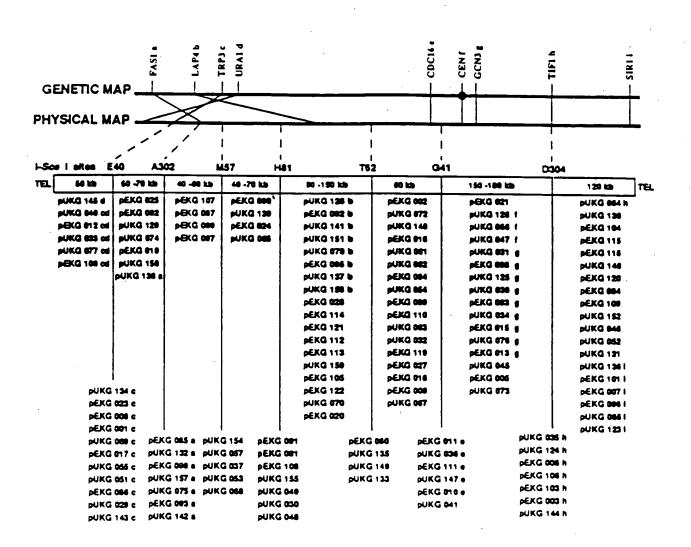


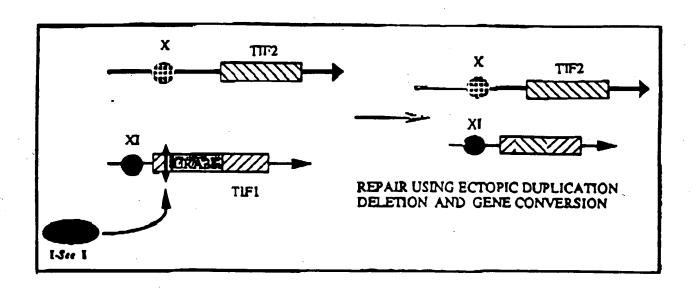


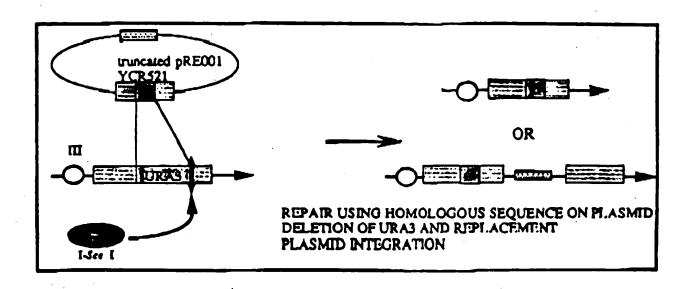


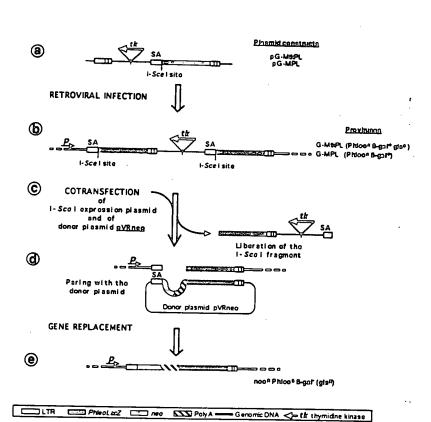


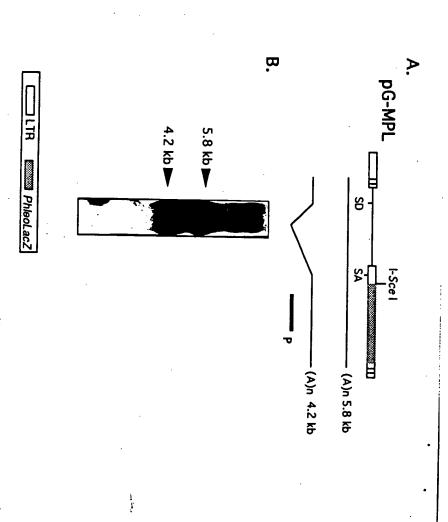


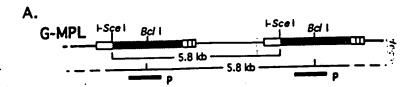


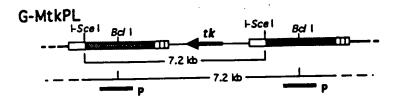








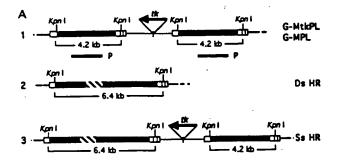


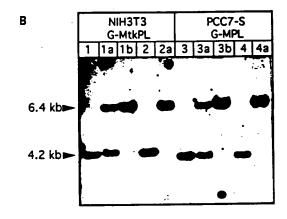


В.

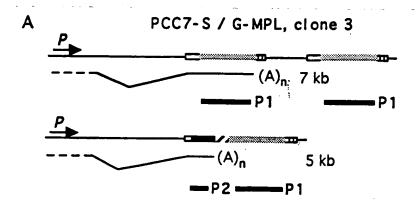
ENZYME		I-Se	:e I			Вс	:/1	
CELL LINE	NIH	3T3	PCC	7-\$		3T3	PCC	
PROVIRUS	G-M	tkPL	G-N	1PL	G-M	tkPL	G-1	4PL
CLONE	1	2	3	4	1_	2	3	4
	ė	·			~~	**	•	
7.2 kb		•			-	•		
5.8 kb -			-			,	•	
						_		•
					."			
							•••	

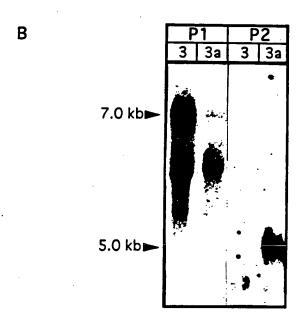
LTR PhleoLacZ tk thymidine kinase

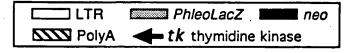




PolyA Genomic DNA tk thymidine kinase







ggers.

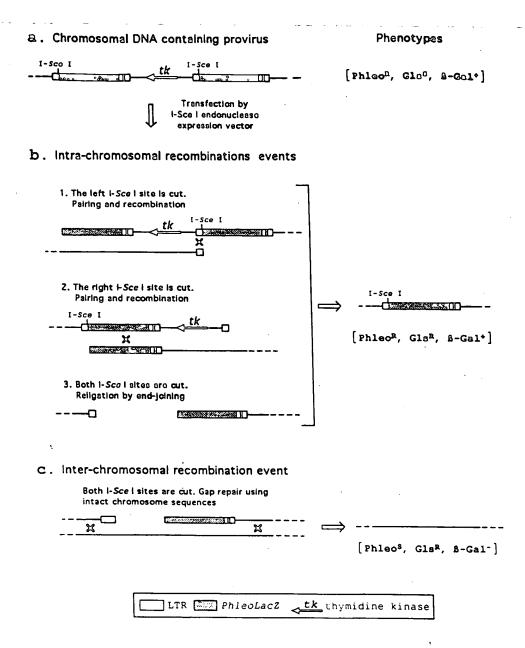
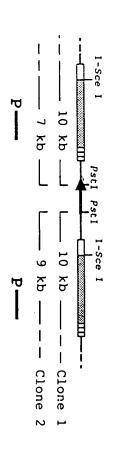


Figure 26

a . Parental DNA, G-MtkPL



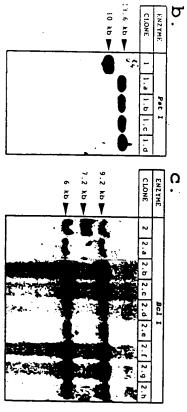
,	CLONE	ENZYME	ტ.
	1 1.1 1.2 1.3 1.4 1.5	Pat I	
9 kb 7 kb V	CLONE	ENZYME	0
	2 2.1 2.2 2.3 2.4 2.5 2.6	Pat I	

2. 1. Parental DNA, G-MtkPL



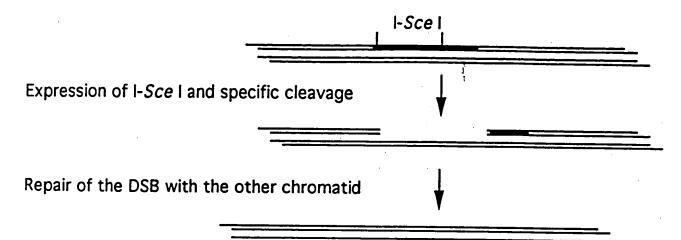
2. Intra-molecular recombination event





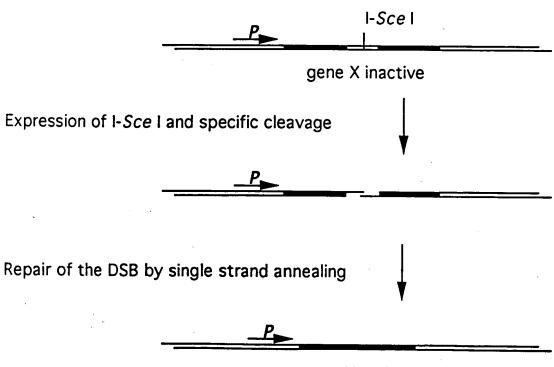
LOSS OF HETEROZYGOSITY

Integration of artificial site or presence of specific site



CONDITIONNAL ACTIVATION (Tandem repeat)

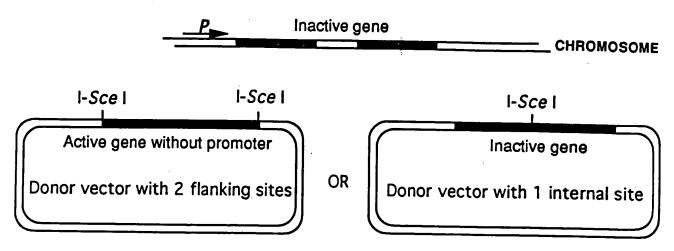
Integration of artificial site between tandem repeats



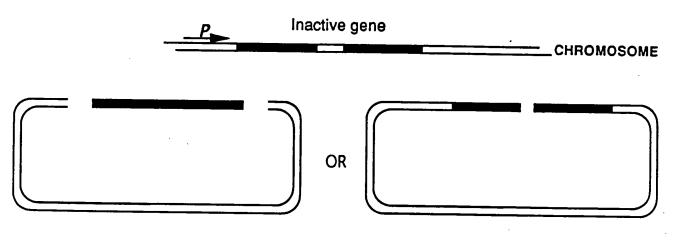
gene X active



Integration of artificial site or presence of specific site



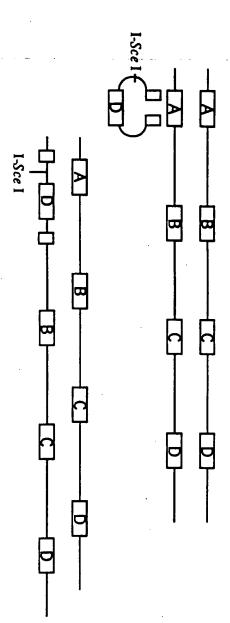
Expression of I-Sce I enzyme and specific cleavage of the donor plasmid



Recombination between the chromosome and plasmid

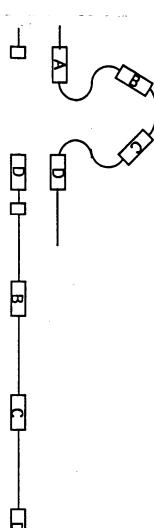


1 insertion of I-Sce I site by classical gene replacement

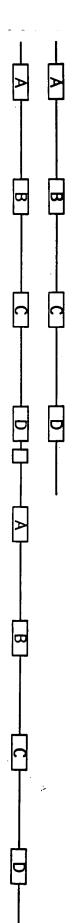


2 Specific cleavage by I-Sce I enzyme and

repair of the break by homologous sequences



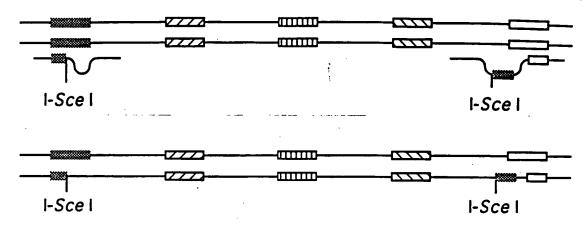
3 Duplication of the totality of the locus



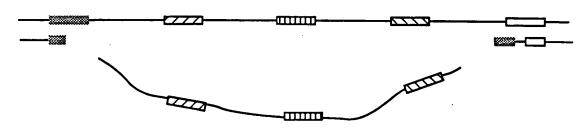




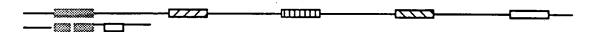
1 Insertion of two I-Sce I sites flanking the locus



2 Expression of the enzyme and cleavage



3 Recombination between the two ends



4 deletion of the locus

